

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 04:41:21 ; Search time 19.3258 seconds  
(without alignments)  
553.709 Million cell updates/sec

Title: US-09-698-781-3

Perfect score: 1436

Sequence: 1 MKQILHPALETMTLFPVL.....KHQVDRSCASCNSNSIY 258

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1363	94.9	245	1	CRS3_HUMAN
2	988	68.8	243	1	TPX1_HUMAN
3	968.5	67.4	244	1	TPX1_CAVPO
4	929	64.7	245	1	CRS3_HORSE
5	830	57.8	243	1	TPX1_MOUSE
6	766.5	53.4	244	1	AEGL_MOUSE
7	755.5	52.6	246	1	AEGL_RAT
8	755.5	52.6	246	1	HELO_HELMO
9	617.5	46.8	241	1	AEGL_MOUSE
10	552	38.4	183	1	CRVP_TRIMU
11	529.5	36.9	249	1	CRS1_HUMAN
12	328	22.8	266	1	GLIF_HUMAN
13	274	19.1	424	1	ASP_ANCCA
14	269	18.7	202	1	VAS1_VESCR
15	265	18.5	202	1	VAS2_VESCR
16	259	18.0	227	1	VAS2_DOLMA
17	257.5	17.9	211	1	VAS3_DOLMA
18	234	17.7	203	1	VAS3_DOLMA
19	232.5	17.6	215	1	VAS3_DOLMA
20	251	17.5	312	1	YR81_CAEEL
21	250.5	17.4	204	1	VAS_VESFL
22	250.5	17.4	205	1	VAS_VESFO
23	250.5	17.4	227	1	VAS_VESFO
24	248	17.3	234	1	VAS3_SOLIN
25	245.5	17.1	204	1	VAS_VESGE
26	243.5	17.0	204	1	VAS_VESMA
27	243	16.9	202	1	VAS_VESMA
28	241.5	16.8	204	1	VAS_VESPE
29	240.5	16.7	206	1	VAS_VESPE
30	228.5	15.9	164	1	PR13_HORVU
31	227.5	15.8	164	1	PR1_HORVU
32	227.5	15.5	206	1	VAS_POLEDO
33	222	15.5	205	1	VAS_POLEX

## ALIGNMENTS

RESULT 1	CRS3_HUMAN	STANDARD:	PRT: 245 AA.
ID	PS4108: 015512:		
AC	PS4108: 015512:		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Cysteine-rich secretory protein-3 precursor (CRISP-3) (SGP28 protein).		
GN	CRISP3.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96270732; PubMed=8665901;		
RA	Kraetzschmar J., Haendler B., Eberspaecher U., Roostermann D.,		
RA	Donner P., Schleuning W.-D.;		
RT	"The human cysteine-rich secretory protein (CRISP) family. Primary		
RT	structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3.";		
RL	Eur. J. Biochem. 236:827-836(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96186934; PubMed=8601434;		
RA	Kjeldsen L., Cowland J.B., Johnson A.H., Borregaard N.;		
RT	"SGP28, a novel matrix glycoprotein in specific granules of human		
RT	neutrophils with similarity to a human testis-specific gene product		
RT	and a rodent sperm-coating glycoprotein.";		
RL	FEBS Lett. 380:246-250(1996).		
CC	-1- SUBCELLULAR LOCATION: SECRETED; IN NEUTROPHILS, LOCALIZED IN		
CC	SPECIFIC GRANULES.		
CC	-1- TISSUE SPECIFICITY: SALIVARY GLAND, PANCREAS AND PROSTATE >		
CC	EPIDIDYMIS, OVARY, THYMUS AND COLON.		
CC	-1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;		
CC	INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by, and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).		
CC	-----		
DR	EMBL: X95240; CAAG4527.1; -		P35792 hordeum vul
DR	EMBL: X94323; CAAG3984.1; -		005109 polistes an
DR	HSSP; P04284; ICPE.		005109 polistes an
DR	InterPro: IPR001283; Allrgn_V5/TPX1.		008697 lycopersico
DR	Pfam: PF00188; SCP. 1.		008697 lycopersico
DR	PRINTS; PR00837; V5PPLIKE.		004108 lycopersico
DR	ProDom; PD000542; Allrgn_V5/TPX1. 1.		041359 sambucus ni
DR	SMART: SM00198; SCP. 1.		P08299 arabidopsi
DR	PROSITE; PS01009; SCP_AG5_PRL_SC7_1; 1.		P33154 arabidopsi
DR	PROSITE; PS01010; SCP_AG5_PRL_SC7_2; 1.		P04280 lycopersico
KW	Glycoprotein; Signal; Multigene family; Polymorphism.		000008 zea mays (m
FT	SIGNAL 1 20 POTENTIAL.		P07053 nicotiana t

```

FT CHAIN 21 245 CYSTEINE-RICH SECRETORY PROTEIN-3.
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 106 106 S -> P (IN DBSNP:495335).
FT VARIANT 134 134 /FTID-VAR.011718.
FT SEQUENCE 245 AA: 27630 MM: B4DD79CB7AE9E5F9 CRC64:
SQ
Query Match 94.9%; Score 1363; DB 1; Length 245;
Best Local Similarity 99.6%; Pred. No. 2.2e-109;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 MTEPLVLLFTVAGLLPSFPANEDKDPFTALLTTQTQVOREIVYKHNELRRVSPPARNM 73
DB 1 MTEPLVLLFTVAGLLPSFPANEDKDPFTALLTTQTQVOREIVYKHNELRRVSPPARNM 60
QY 74 LKHEMNKEAANAKMANOCNRYRHSNPKDRMTSLKCGENLYMSSAPSSWSOAIOSWFDEY 133
DB 61 LKHEMNKEAANAKMANOCNRYRHSNPKDRMTSLKCGENLYMSSAPSSWSOAIOSWFDEY 120
QY 134 NDEDFGVGPKTPNNAVGHYTVQVWYSSYLVCGNAYCPNOKVLKYYVCOYCPAGNMNMR 193
DB 121 NDEDFGVGPKTPNNAVGHYTVQVWYSSYLVCGNAYCPNOKVLKYYVCOYCPAGNMNMR 180
QY 194 LYPYEGGAPCASPDCDGLCTNGCKYEDLYSNCKSLKLTTLCKHQLYRDSCKASCNC 253
DB 181 LYPYEGGAPCASPDCDGLCTNGCKYEDLYSNCKSLKLTTLCKHQLYRDSCKASCNC 240
QY 254 SNSIY 258
DB 241 SNSIY 245

RESULT 2
TPX1_HUMAN STANDARD; PRT; 243 AA.
ID TPX1_HUMAN
AC P16562;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Testis-specific protein TPX-1 precursor (Cysteine-rich secretory
DE protein-2) (Crisp-2).
GN TPX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RA MEDLINE=96270732; PubMed=8665901;
RA Kraetzschmar J., Haendler B., Eberspaecher U., Roostermann D.,
RA Donner P., Schleuning W.D.;
RT "The human cysteine-rich secretory protein (CRISP) family. Primary
RT structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3.";
RL Eur. J. Biochem. 236:827-836(1996).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: TESTIS AND EPIDIDYMIS.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AGS; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25532; AAA61220.1; -
DR EMBL: X95239; CAA64526.1; -
DR PIR: B33329; B33329.
DR Genew; HGNC:12024; TPX1.
DR MIM; 187430; -
DR InterPro; IPR001283; Allrgn_V5/tpx1.
DR Pfam; PF00188; SCP_1.
DR PRINTS; PR00837; VSTPXLIKE.
DR ProDom; PD000542; Allrgn_V5/tpx1; 1.
DR SMART; SM00198; SCP_1.
DR PROSITE; PS01009; SCP_AGS_PRT_SG7_1; 1.
DR PROSITE; PS01010; SCP_AGS_PRT_SG7_2; 1.
KW Testis; signal; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 243
SQ SEQUENCE 243 AA: 27259 MM: C5FE698C449CFAD9 CRC64:
SQ
Query Match 68.8%; Score 988; DB 1; Length 243;
Best Local Similarity 71.8%; Pred. No. 2.2e-77;
Matches 176; Conservative 25; Mismatches 42; Indels 2; Gaps 2;

QY 14 MTEPLVLLFTVAGLLPSFPANEDKDPFTALLTTQTQVOREIVYKHNELRRVSPPARNM 73
DB 1 MALLPV-LFLVTVLPSLPA-EGKDPFTALLTTQTQVOREIVYKHNELRRVSPPARNM 58
QY 74 LKHEMNKEAANAKMANOCNRYRHSNPKDRMTSLKCGENLYMSSAPSSWSOAIOSWFDEY 133
DB 59 LKHEMSREVTNMGARMANKCTLOHSDPEDKRTSRCEENLYMSDPSWSOAIOSWFDEI 118
QY 134 NDEDFGVGPKTPNNAVGHYTVQVWYSSYLVCGNAYCPNOKVLKYYVCOYCPAGNMNMR 193
DB 119 LDFYGVGPKSPNNAVGHYTVQVWYSSYLVCGNAYCPNODSLKYYVCOYCPAGNMNMR 178
QY 194 LYPYEGGAPCASPDCDGLCTNGCKYEDLYSNCKSLKLTTLCKHQLYRDSCKASCNC 253
DB 179 KNTFYQGCTGACGPDGCDGLCTNSCYODLNSCDLSLKNVAGCEHLLKCKKATCCLC 238
QY 254 SNSIY 258
DB 239 ENKIY 243

RESULT 3
TPX1_CAVPO STANDARD; PRT; 244 AA.
ID TPX1_CAVPO
AC O60477;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Testis-specific protein TPX-1 precursor (Autoantigen 1) (25 kDa
DE acrosomal autoantigen) (AA1).
GN TPX1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley; TISSUE=Testis;
RA MEDLINE=96354287; PubMed=9115720;
RA Foster J.A., Gerton G.L.;
RT "Autoantigen 1 of the guinea pig sperm acrosome is the homologue of
RT mouse Tpx-1 and human TPX1 and is a member of the cysteine-rich
RT secretory protein (CRISP) family.";
RL Mol. Reprod. Dev. 44:221-229(1996).
RN [2]
RP SEQUENCE OF 22-41.
RA MEDLINE=88193219; PubMed=3282555;
RA Hardy D.M., Huang T.T.F. Jr., Driscoll W.J., Tung K.S.K., Wild G.C.;

```

RT "Purification and characterization of the primary acrosomal  
 RT autoantigen of guinea pig epididymal spermatozoa."  
 RL Biol. Reprod. 38:423-437(1988).  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U35712; AAC52616.1; -  
 DR InterPro: IPR001283; Allrgn\_V5/TPx1.  
 DR Pfam: PF00188; SCP; 1.  
 DR PRINTS: PR00837; V5TPXLIKE.  
 DR ProDom: PD000542; Allrgn\_V5/TPx1; 1.  
 DR SMART: SM00198; SCP; 1.  
 DR PROSITE: PS01009; SCP\_AGS\_PRL\_SC7\_1; 1.  
 DR PROSITE: PS01010; SCP\_AGS\_PRL\_SC7\_2; 1.  
 DR Testis: Signal.  
 KW Testis: Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21  
 SQ SEQUENCE 244 AA: 27248 MW: 5808DE6EECE1A12 CRC64;  
 -----  
 Query Match 67.4%; Score 968.5; DB 1; Length 244;  
 Best Local Similarity 70.2%; Pred. No. 1e-75;  
 Matches 172; Conservative 26; Mismatches 46; Indels 1; Gaps 1;  
 QY 14 MTLFVLLFLVAGLPSFPANEDKDPATLALTTOVOGREIYNKNELRRAVSPARRM 73  
 DB 1 MALLPVLELITMLTLPCLVLTN-GKDPATLALTTOVOONEIINKHNOIARKSTPPASNM 59  
 QY 74 LKEMNKKEAANAOKKAWANOCNRRHSNPKDRMTSLKCGENLYSSAPSSMSQAIOSFDE 133  
 DB 60 LKEMNSREAAVNAOKKAWANCTLVHSNPDKRSTKCGENLYSSDPSSMSDAIOSFDES 119  
 QY 134 NDFEGVGPKTPNAVVGHTOVVWYSSYLVGCGNAYCPNOKYLKYYVQYCPAGMANM 193  
 DB 120 QPFTGVGPKSHNAVVGHTOLVWYSSYLVGCGIAYCPNODSLKYYVQYCPAGNNVT 179  
 QY 194 RLVPPYEGAPCASPDCNDGDLCTNGCKYEDLYSNCKSLKLTLTCKHOLVRDSCASCNC 253  
 DB 180 KMTPEYQGTPCARCPGNCNDGDLCTNSCEYEDLYSNCKSLKLTAGCEHOLVREKCRATC 239  
 QY 254 SNSIT 258  
 DB 240 EDKIT 244  
 -----  
 RESULT 4  
 CRS3 HORSE  
 ID CRS3 HORSE STANDARD; PRT: 245 AA.  
 AC 019010;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cysteine-rich secretory protein-3 precursor (CRISP-3).  
 GN CRISP3.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 OX [1]  
 RN RP  
 RC TISSUE=Amplula;  
 RC MEDLINE=98422318; PubMed=9748582;  
 RA Schambony A., Gentzel M., Wolfes H., Raida M., Neumann U.,  
 RA Toepfer-Petersen E.;  
 RA \*Equine CRISP-3: primary structure and expression in the male genital

RT tract."; Biophys. Acta 1387:206-216(1998).  
 RL Biochim. Biophys. Acta 1387:206-216(1998).  
 CC -1- SUBCELLULAR LOCATION: SECRETED; IN NEUTROPHILS, LOCALIZED IN  
 CC SPECIFIC GRANULES (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE SALIVARY GLAND, IN THE  
 CC AMPULLA AND THE SEMINAL VESICLE.  
 CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ001400; CA04729.1; -  
 DR HSP: P04284; ICEE.  
 DR InterPro: IPR001283; Allrgn\_V5/TPx1.  
 DR Pfam: PF00188; SCP; 1.  
 DR PRINTS: PR00837; V5TPXLIKE.  
 DR ProDom: PD000542; Allrgn\_V5/TPx1; 1.  
 DR SMART: SM00198; SCP; 1.  
 DR PROSITE: PS01009; SCP\_AGS\_PRL\_SC7\_1; 1.  
 DR PROSITE: PS01010; SCP\_AGS\_PRL\_SC7\_2; 1.  
 DR Signal: Multigene family.  
 KW Signal: Multigene family.  
 FT SIGNAL 1 22  
 FT CHAIN 1 22  
 SQ SEQUENCE 245 AA: 27308 MW: 893A4E87F402BA22 CRC64;  
 -----  
 Query Match 64.7%; Score 929; DB 1; Length 245;  
 Best Local Similarity 65.9%; Pred. No. 2.4e-72;  
 Matches 162; Conservative 33; Mismatches 49; Indels 2; Gaps 2;  
 QY 14 MTLFVLLFLVAGLPSFPANEDKDPATLALTTOVOGREIYNKNELRRAVSPARRM 73  
 DB 1 MALLPVLELITMLTLPCLVLTN-GKDPATLALTSTKSEVOKEIYNKHNDRRTVPLASNM 59  
 QY 74 LKEMNKKEAANAOKKAWANOCNRRHSNPKDR-MTSLKCGENLYSSAPSSMSQAIOSFDE 132  
 DB 60 LKQMDSKTATNANNANNCILQHSKAEDPRAVGTMKCGENLFMSIPNSWSDAIQWHE 119  
 QY 133 YNDFEGVGPKTPNAVVGHTOVVWYSSYLVGCGNAYCPNOKYLKYYVQYCPAGMANM 192  
 DB 120 VDFEKGVGPKTPNAVVGHTOVVWYSSYRVGCGIAYCPNOKYLKYYVQYCPAGMANM 179  
 QY 193 RLVPPYEGAPCASPDCNDGDLCTNGCKYEDLYSNCKSLKLTLTCKHOLVRDSCASCNC 252  
 DB 180 KMTPEYQGTPCARCPGNCNDGDLCTNSCEYEDLYSNCKSLKLTAGCEHOLVREKCRATC 239  
 QY 253 CSNSIT 258  
 DB 240 CENKIT 245  
 -----  
 RESULT 5  
 TPX1 MOUSE  
 ID TPX1 MOUSE STANDARD; PRT: 243 AA.  
 AC P16563;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Testis-specific protein TPX-1 precursor.  
 GN TPX1 OR TPX-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 OX [1]  
 RN RP  
 RC TISSUE=Testis;  
 RC MEDLINE=90129048; PubMed=2613236;

RA Kasahara M., Gutknecht J., Brew K., Spurr N., Goodfellow P.N.;  
 RT "Cloning and mapping of a testis-specific gene with sequence  
 RL similarity to a sperm-coating glycoprotein gene.";  
 CC Genomics 5:527-534(1989).  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC  
 CC EMBL: M25533; AAA0472.1; .  
 CC PIR: A33329; A33329.  
 CC HSSP: P04284; ICPE.  
 CC MGD: MGI:98815; Tpx1.  
 CC InterPro: IPR001283; Allrgn\_V5/TPx1.  
 CC Pfam: PF00188; SCP; 1.  
 CC PRINTS: PR00837; V5TPXLIKE.  
 CC PRODOM: PD000542; Allrgn\_V5/TPx1; 1.  
 CC SMART: SM00198; SCP; 1.  
 CC PROSITE: PS01009; SCP\_AG5\_PRL\_SCP\_1; 1.  
 CC PROSITE: PS01010; SCP\_AG5\_PRL\_SCP\_2; 1.  
 CC Testis; Signal.  
 CC SIGNAL 1 22 POTENTIAL.  
 CC CHAIN 23 243 TESTIS-SPECIFIC PROTEIN TPX-1.  
 CC SEQUENCE 243 AA; 27605 MM; 6E707F569ACAA244 CMC64;  
 SO  
 Query Match 57.8%; Score 830; DB 1; Length 243;  
 Best Local Similarity 59.6%; Pred. No. 6.8e-64;  
 Matches 146; Conservative 39; Mismatches 58; Indels 2; Gaps 2;  
 QY 14 MTLFVLLFLVAGLSPFANEDKDPATALLTTOTVORELYVKNHLELRRAVSPAPNM 73  
 DB 1 MAMTQVLFVFAALLRLS-PLTEGKDPDTSLITNOLQVORELYVKNHLELRRAVSPAPNM 59  
 QY 74 LKMEWKEAANAOKMAQNCYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIOQMFDEY 133  
 DB 60 LKMEWSIQTATNAOKMAKCIHSHSSKDRKINIRCGENLYMSTPTLMSTYIOGWYEN 119  
 QY 134 NDFDGVGPKTPNAVGHYTVVWYSSYLVCGNAYCPNOKLYKYVCQCPAGNMNR 193  
 DB 120 EDFYVGVGAK-PNSAVGHYTVVWYSSFKIGCIAYCPNODLKIFYVCHYCPMGNNMVK 178  
 QY 194 LYPYEGAGPASCPCDNCDDGLCTNGCKYEDLYSNCKSLKLTLCCKHQLVBDSCASCNC 253  
 DB 179 KSTRPGQSTPCASCPNCKENGLCTNSCDFEILSNCKESLTKTSAGCKHLLTKCQATLC 238  
 QY 254 SNSIY 258  
 DB 239 EDKIH 243

RP SEQUENCE FROM N.A.  
 RC TISSUE-Submandibular gland;  
 RA MEDLINE-93346016; PubMed-1301383;  
 RX Mizuki N., Kasahara M.;  
 RT "Mouse submandibular glands express an androgen-regulated transcript  
 RL encoding an acidic epididymal glycoprotein-like molecule.";  
 RN Mol. Cell. Endocrinol. 89:25-32(1992).  
 RP  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Submandibular gland;  
 RX MEDLINE-93307144; PubMed-8319566;  
 RA Haendler B., Kratzschmar J., Theuring F., Schleuning M.D.;  
 RT "Transcripts for cysteine-rich secretory protein-1 (CRISP-1; DE/AGC)  
 RT and the novel related CRISP-3 are expressed under androgen control in  
 RT the mouse salivary gland.";  
 RL Endocrinology 133:192-198(1993).  
 CC  
 CC -1- FUNCTION: THIS PROTEIN IS SUPPOSED TO HELP SPERMATOZOA UNDERGO  
 CC FUNCTIONAL MATURATION WHILE THEY MOVE FROM THE TESTIS TO THE  
 CC DUCTUS DEFERENS.  
 CC  
 CC -1- SUBCELLULAR LOCATION: STORED IN SECRETORY GRANULES OF GRANULAR  
 CC CONVOLUTED TUBULES CELLS.  
 CC  
 CC -1- TISSUE SPECIFICITY: MAINLY FOUND IN THE CAUDA EPIDIDYMIS WHERE IT  
 CC IS SYNTHESIZED BY THE PRINCIPAL CELLS AND SECRETED INTO THE LUMEN.  
 CC BINDS TO THE HEADS OF SPERMATOZOON. ALSO EXPRESSED IN THE  
 CC SUBMANDIBULAR GLAND.  
 CC  
 CC -1- DEVELOPMENTAL STAGE: EXPONENTIAL INCREASE BETWEEN DAYS 25 AND 30  
 CC AFTER BIRTH.  
 CC  
 CC -1- INDUCTION: By androgens.  
 CC  
 CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;  
 CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC  
 CC EMBL: M92849; AAA37185.1; .  
 CC EMBL: L05559; AAA37460.1; .  
 CC PIR: A49202; AAA37460.1; .  
 CC MGD: MGI:102553; Aeg1.  
 CC InterPro: IPR001283; Allrgn\_V5/TPx1.  
 CC Pfam: PF00188; SCP; 1.  
 CC PRINTS: PR00837; V5TPXLIKE.  
 CC PRODOM: PD000542; Allrgn\_V5/TPx1; 1.  
 CC SMART: SM00198; SCP; 1.  
 CC PROSITE: PS01009; SCP\_AG5\_PRL\_SCP\_1; 1.  
 CC PROSITE: PS01010; SCP\_AG5\_PRL\_SCP\_2; 1.  
 CC Sperm; Glycoprotein; Signal.  
 CC SIGNAL 1 19  
 CC CHAIN 20 244 SPERM-COATING GLYCOPROTEIN 1.  
 CC CARBOHYD 145 145 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC SEQUENCE 244 AA; 27679 MM; D00DD0348F85781F CMC64;  
 SO  
 Query Match 53.4%; Score 766.5; DB 1; Length 244;  
 Best Local Similarity 55.1%; Pred. No. 1.8e-58;  
 Matches 135; Conservative 39; Mismatches 70; Indels 1; Gaps 1;  
 QY 14 MTLFVLLFLVAGLSPFANEDKDPATALLTTOTVORELYVKNHLELRRAVSPAPNM 73  
 DB 1 MAMTQVLFVFAALLRLS-PLTEGKDPDTSLITNOLQVORELYVKNHLELRRAVSPAPNM 59  
 QY 74 LKMEWKEAANAOKMAQNCYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIOQMFDEY 133  
 DB 61 LKMEWKEAANAOKMAQNCYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIOQMFDEY 120  
 QY 134 NDFDGVGPKTPNAVGHYTVVWYSSYLVCGNAYCPNOKLYKYVCQCPAGNMNR 193  
 DB 121 KDLVYDVGPKPDVSVGHYTVVWYSSYLVCGNAYCPNOKLYKYVCQCPAGNMNR 179  
 QY 194 LYPYEGAGPASCPCDNCDDGLCTNGCKYEDLYSNCKSLKLTLCCKHQLVBDSCASCNC 253

```

Db 180 LPTYPAGEPCASCPDHCEDGLCTNSCGHEDKYTNCKTKLKKMLSCSEHELLKGCATGCLC 239
OY 254 SNSIT 258
Db 240 EGRKH 244

RESULT 7
AEG_RAT
ID AEG_RAT STANDARD: PRT: 246 AA.
AC P12020;
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sperm-coating glycoprotein precursor (SCP) (Acidic epididymal
glycoprotein) (Protein D) (Protein E) (Protein IV) (Staloprotein)
DE (32 kDa epididymal protein).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-Epididymis;
RX MEDLINE=87053995; PubMed=3780731;
RA Brooks D.E., Means A.R., Wright E.J., Singh S.P., Tiver K.K.;
RT "Molecular cloning of the cDNA for androgen-dependent sperm-coating
glycoproteins secreted by the rat epididymis."
RL Eur. J. Biochem. 161:13-18(1986).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=89039913; PubMed=2460753;
RA Charrest N.J., Joseph D.R., Wilson E.M., French F.S.;
RT "Molecular cloning of complementary deoxyribonucleic acid for an
androgen-regulated epididymal protein: sequence homology with
metalloproteins."
RL Mol. Endocrinol. 2:999-1004(1988).
CC -1- FUNCTIONAL: THIS PROTEIN IS SUPPOSED TO HELP SPERMATOZOEA UNDERGO
FUNCTIONAL MATURATION WHILE THEY MOVE FROM THE TESTIS TO THE
DUCTUS DEFERENS.
CC -1- INDUCTION: By androgens.
CC -1- MISCELLANEOUS: SCP IS AN ANDROGEN-DEPENDENT PROTEIN, WHICH IS
SECRETED BY THE EPIDIDYMAL EPITHELIUM AND THEN BECOMES ASSOCIATED
WITH THE SPERM SURFACE. TWO MAJOR VARIANT PROTEIN D AND E DIFFER
FROM EACH OTHER BY THEIR CARBOHYDRATE SIDE CHAINS.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X04643; CAA28304.1;
DR EMBL: M31173; AAB59716.1;
DR PIR: A24609; A24609.
DR PIR: A40918; A40918.
DR InterPro: IPR001283; Allrgn_V5/TPX1.
DR Pfam: PF00188; SCP_1.
DR PRINTS: PRO0837; V5TPXLIKE.
DR PRODOM: PD000542; Allrgn_V5/TPX1; 1.
DR SMART: SM00198; SCP_1.
DR PROSITE: PS01009; SCP_AG5_PRL_SC7_1; 1.
DR PROSITE: PS01010; SCP_AG5_PRL_SC7_2; 1.
KW Sperm; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 246 SPERM-COATING GLYCOPROTEIN.
FT MOD_RES 20 20 BLOCKED (POTENTIAL).
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 246 AA: 27847 MW; 585207C7CF7CE9D1 CRC64.

Query Match 52.6%; Score 755.5; DB 1; Length 246;
Best Local Similarity 55.1%; Pred. No. 1.6e-57;
Matches 136; Conservative 35; Mismatches 73; Indels 3; Gaps 2;

OY 14 MTLFPVLLFLVAGLPSF--PANEDKPAFTALLTTOTVOGREIYNKHNELRANVSPAR 71
Db 1 MALMLVLLFLAAVLPSPSLQDITDEMIDRLNLTSTKLSEVDEIIINKHQLRRTSPSGS 60
OY 72 NNLKMKMKKKAANOKNANOCNRYHSNPKDMTSLKGCENLYMSAPSMSQAIOQSPFD 131
Db 61 DLRLVEMHDIAVYNAOKNANRCITYNHSPLOHHTTLTKCGENLFMANYPASMSVYIQDWID 120
OY 132 EYNDEDFGVGPKTPNAVVGHYTOVVMYSSYLVCGNAYCPNOKVLYKYYVCOCPAGNMA 191
Db 121 ESLDFVFGFGPKKVGKGYHYTOVVMNSTFLVACGVAECPDGP-LKTFYVCHYCGGNVY 179
OY 192 NRLVYPTGEGAPCASCPCPDNDGLCTNGCKYEDLYSNCKSLKLLITCKHQLVRSCKASC 251
Db 180 GLTSPYTEGEPDCDCPCNCEDEGLCTNSCEYEDNYSNCGDLKKMYSCDPLLEKGCRA SC 239
OY 252 NCSNSIT 258
Db 240 FLEDKIH 246

RESULT 8
HELO_HELHO
ID HELO_HELHO STANDARD: PRT: 242 AA.
AC 091055;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Helothermine precursor (HLTX).
OC Heloderma horridum horridum (Mexican beaded lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguilliformes; Helodermatidae;
OC Heloderma.
OX NCBI_TaxID=8552;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-Salivary gland;
RX MEDLINE=95375162; PubMed=7647234;
RA Morrissette J., Kraetzschmar J., Haendler B., El-Hayek R.,
RA Mochica-Morales J., Martin B.M., Patel J.R., Moss R.L.,
RA Schenking W.-D., Coronado R., Poesani L.D.;
RT "Primary structure and properties of helothermine, a peptide toxin
that blocks ryanodine receptors."
RL Biophys. J. 68:2280-2288(1995).
RN (2)
RP SEQUENCE OF 20-39, AND CHARACTERIZATION.
RX TISSUE-Venom;
RX MEDLINE=90260878; PubMed=1693019;
RA Mochica-Morales J., Martin B.M., Poesani L.D.;
RT "Isolation and characterization of helothermine, a novel toxin from
Heloderma horridum horridum (Mexican beaded lizard) venom."
RL Toxicon 28:299-309(1990).
CC -1- FUNCTION: Toxin that blocks ryanodine receptors. It is toxic to
mice: it causes lethargy, partial paralysis of rear limbs and
lowering of body temperature.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Produced by the venomous gland.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

```

```

CC - FUNCTION: THIS PROTEIN IS SUPPOSED TO HELP SPERMATOZOA UNDERGO
CC FUNCTIONAL MATURATION WHILE THEY MOVE FROM THE TESTIS TO THE
CC DUCTUS DEFERENS.
CC - SUBCELLULAR LOCATION: STORED IN SECRETORY GRANULES OF GRANULAR
CC CONVOLUTED TUBULES CELLS.
CC - TISSUE SPECIFICITY: EXPRESSED IN SUBMANDIBULAR GLAND.
CC - DEVELOPMENTAL STAGE: EXPONENTIAL INCREASE BETWEEN DAYS 25 AND 30
CC AFTER BIRTH.
CC - INDUCTION: By androgens.
CC - SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -----
CC This Swiss-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/).
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M92850; AAA37186.1; -
CC EMBL: L05560; AAA37461.1; -
CC PIR: B49202; B49202.
CC MGD: MGI:102552; Aeg2.
CC InterPro: IPR001283; Allrgn_V5/Trpx1.
CC Pfam: PF00188; SCP; 1.
CC PRINTS: PR00837; V5TPRLIKE.
CC PRODOM: PR000542; Allrgn_V5/Trpx1; 1.
CC SMART: SM00198; SCP; 1.
CC PROSITE: PS01009; SCP_AG5_PRL_SC7_1; 1.
CC DR PROSITE: PS01010; SCP_AG5_PRL_SC7_2; 1.
CC KW Sperm; Glycoprotein; Signal.
CC FT SIGNAL 1 19
CC CHAIN 20 241 SPERM-COATING GLYCOPROTEIN 2.
CC CARBOHD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT SEQUENCE 241 AA; 27314 MM; D903788BA4E4001EF CRC64;
SO
Query Match 43.0%; Score 617.5; DB 1; Length 241;
Best Local Similarity 48.4%; Pred. No. 9.1e-46;
Matches 118; Conservative 40; Mismatches 79; Indels 7; Gaps 3;
OY 14 MTLPEVLFLVAGLLSPSPANEKDPATFALLTTOTQVQRELVNNHNLRAVSPARNM 73
1 | | | | | | | | | | : : : : | | | | | | | | : :
DB 1 MALMLVLFLLAVALRPSSLIDONSQNSLEKISTSKSVQEEIVSHNQDLRRKVPSSGDL 60
74 LKMEENKRAANAQKMANQCNHRHSNPRDKMTSLKCGENLYSSAPSSQSAIQSFDPEY 133
1 | | | | | | | | | | : : : : | | | | | | | | : :
DB 61 LNMENNYDAQVNAQDRAKCFTHSHPIELRTTNLKGLENLFMSVLYQWMSVYIDGWYNES 120
134 NDFDGVGPKTPMAVAVGHYTOVYVWSSVLYVGGNAYFCNOUKLKYUYCOXCP----AGN 189
| | | | | | | | | | : : | | | | | | | | : :
DB 121 KGLIGVGPRKQNVSVYGHHTQVWKMNSNQVACGVAECF-ENPLKRFYFCRCRPVLNYSGH 179
OY 190 WANRLVYVEDGAPCARPCDNDGLCTNGCKIEDLYSNCSSLKLTTLTCKHQLVNDSCKA 249
: : : : | | | | | | | | | | : : | | | | : :
DB 180 YPSRFLVLYTAARAPCASCPRDCEBDGLCTKSCQYKDMSEFWCK--RLEVYCKHGPLKRCCLA 237
OY 250 SCNC 253
: : |
DB 238 TCQC 241

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidostei; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Protobothrops.
OX NCBI_TaxID=103944;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE-Venom gland;
RX MEDLINE=97385768; PubMed=9241782;
RA Chang T.-Y., Mao S.-H., Guo Y.-W.;
RT Cloning and expression of a cysteine-rich venom protein from
RT Toxicon 35:879-888(1997).
RL Toxicon 35:879-888(1997).
CC -1- FUNCTION: May be a toxin that blocks ryanodine receptors.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Produced by the venomous gland.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U59447; AAB48565.1; -
DR InterPro: IPR001283; Allrgn_V5/TPx1.
DR Pfam: PF00188; SCP; 1.
DR PRINTS: PR00837; V5TPXLIKE.
DR ProDom: PD000542; Allrgn_V5/TPx1; 1.
DR SMART: SM00198; SCP; 1.
DR PROSITE: PS01009; SCP_AGS_PRL_SC7_1; 1.
DR PROSITE: PS01010; SCP_AGS_PRL_SC7_2; 1.
DR Calcium channel inhibitor; Neurotoxin; Toxin; Signal.
KM SIGNAL 1 21
FT CHAIN 1 21
SQ SEQUENCE 183 AA; 20378 MW; 50224DBE92A64930 CRC64;

Query Match
Best Local Similarity 51.4%; Score 553; DB 1; Length 183;
Matches 94; Conservative 23; Mismatches 64; Indels 2; Gaps 2;

OY 76 MEMKKAANAOQKMANOCVYRHSNPKDRMT-SLKCGENTLYMSSAPSSQAIOSWFEYV 134
DB 1 MEMYPPAANAERMAIRCRLESHSRSDSRVIGCKGCEINIMSPYAKWIDIHANMGEYK 60
OY 135 DFDFGVGPTPPNAVVGHTYQVWYSSYLVCGNAYCPNOKVLYKYVYVCOYCPAGNNANRL 194
DB 61 DERKYGAVPSPNAAHTHYQIYWYKSRGGCAAAAYCPSSK-YRYFYVCOYCPAGNNIGKT 119
OY 195 YVPEYEGACAPDCNDGCLCTNGCKYEDLYSNCKSLKTLTKCKQLVYDSCKASCNC 254
DB 120 APTPTSGPPCGDCPCDNDGCLCTNCTOETWTSNCSNLVQSSCQDNNMKTCPASCFCQ 179
OY 255 NSI 257
DB 180 NKI 182

RESULT 11
CRSL_HUMAN STANDARD; PRT; 249 AA.
AC PS4107; 013248; 000698; 014082;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteine-rich secretory protein-1 precursor (Acidic epididymal
DE glycoprotein homolog) (AEG-like protein) (ARP).
GN AEG1 OR CRISP-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), AND SEQUENCE OF 22-41.
RX MEDLINE=96270732; PubMed=8665901;
RA Kretzschmar J., Haendler B., Eberspaecher U., Roostermann D.,
RA Donner P., Schleuning W.-D.;
RT "The human cysteine-rich secretory protein (CRISP) family. Primary
RT structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3."
RT Eur. J. Biochem. 236:827-836(1996).
RN (2)
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RX MEDLINE=96103955; PubMed=8543280;
RA Hayashi M.;
RT "Analysis of the human acidic epididymal glycoprotein-like molecule:
RT isolation of cDNA and tissue localization."
RT Hokkaido Igaku Zasshi 70:743-753(1995).
RN (3)
RP SEQUENCE FROM N.A. (LONG ISOFORM), AND CHARACTERIZATION.
RX MEDLINE=96435914; PubMed=8838800;
RA Hayashi M., Fujimoto S., Takano H., Ushiki T., Abe K., Ishikura H.,
RA Yoshida M.C., Kirchhoff C., Ishibashi T., Kasahara M.;
RT "Characterization of a human glycoprotein with a potential role in
RT sperm-egg fusion: cDNA cloning, immunohistochemical localization,
RT and chromosomal assignment of the gene (AEG1)."
RL Genomics 32:367-374(1996).
CC -1- FUNCTION: MAY HAVE A ROLE IN SPERM-EGG FUSION AND MATURATION.
CC -1- SUBCELLULAR LOCATION: LOCATED IN THE LUMEN AND EPITHELIUM OF
CC DISTAL DUCTUS EPIDIDYMUS AND EPIDIDYMAL DUCTS, AND ON THE
CC POSTACROSOMAL REGION OF THE SPERM HEAD.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form/CRISP-1 delta; may be produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: CAPUT, CORPUS, AND CAUDA REGIONS OF THE
CC EPIDIDYMUS. THE DUCTUS DEFERENS, SPERM, AND SEMINAL PLASMA.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X95237; CA64524.1; -
DR EMBL: X95238; CA64525.1; -
DR EMBL: 580310; AAB35889.1; -
DR EMBL: D38451; BAA07483.1; -
DR GeneW: HGNC:304; AEG1.
DR MIM: 601193; -
DR InterPro: IPR001283; Allrgn_V5/TPx1.
DR Pfam: PF00188; SCP; 1.
DR PRINTS: PR00837; V5TPXLIKE.
DR ProDom: PD000542; Allrgn_V5/TPx1; 1.
DR SMART: SM00198; SCP; 1.
DR PROSITE: PS01009; SCP_AGS_PRL_SC7_1; 1.
DR PROSITE: PS01010; SCP_AGS_PRL_SC7_2; 1.
DR Glycoprotein; Signal; Multigene family; Alternative splicing.
FT SIGNAL 1 21
FT CHAIN 1 21
FT CARBOHYD 22 249
FT VARSPLIC 178 178
FT VARSPLIC 179 249
FT CONFLICT 22 22
FT CONFLICT 97 97
SQ SEQUENCE 249 AA; 28481 MW; 85BED3DEAD62A9C CRC64;

Query Match
Best Local Similarity 36.9%; Score 529.5; DB 1; Length 249;
Matches 111; Conservative 26; Mismatches 100; Indels 7; Gaps 5;

OY 20 LFLVAG--LLP--SEPNADKDPATALLTGTQVREIVNKHNEIRRAVSPARRMLK 75
||||| ||| | : | : | |||| |||| | ||| |||

```

```

DB 6 LFLVAAACLPMLSMKKKSARD-QFNKLVTDLPNVOEYININALRRVPPASNMK 64
QY 76 MEANKEAANQKWAQNDNYRHSNPKD-RMTSLKCGENTLYNSAPSSWSQAIOGFDEYN 134
DB 65 MSWSEEAQNARISKSCDMTESNPLERLPNFCGEMNMTSPVSMSSVIGWYSSST 124
QY 135 DFDGVEGKRTNAV-VGHTYOVWYSSYLVCAGNAYCNOVLYLKYIYQYCPAGNMNR 193
DB 125 SFKGEWTTTDDITTDHYTOYVATSYLCAIASCROGSPRLYLVCYCHCEGNDPET 184
QY 194 LYVPEOGAPCASPDCNDGCTNGCKEYEDLYSNCKSLKLTLCCKHQLVADSCASCNC 253
DB 185 KNEPKYGVPCDAGCSNCEKLTNPCTIYDEYFDGDIQVHYLCCNHSITLFCATCLC 244
QY 254 SNSI 257
DB 245 DTEI 248

RESULT 12
GLIP_HUMAN
ID GLIP_HUMAN STANDARD: PRT: 266 AA.
AC PA8060: 015409:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glioma pathogenesis-related protein (GLIPR) (RTVP-1 protein).
GN GLIPR OR RTVP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Glial tumor:
RX MEDLINE=97128816; PubMed=8973356;
RA Rich T., Chen P., Furman F., Huynh N., Israel M.A.;
RT "RTVP-1, a novel human gene with sequence similarity to genes of
RT diverse species, is expressed in tumor cell lines of glial but not
RT neuronal origin.";
RL Gene 180:125-130(1996).
RN [2]
RP SEQUENCE OF 11-266 FROM N.A.
RX MEDLINE=95331646; PubMed=7607567;
RA Murphy E.V., Zhang Y., Zhu W., Biggs J.;
RT "The human glioma pathogenesis-related protein is structurally
RT related to plant pathogenesis-related proteins and its gene is
RT expressed specifically in brain tumors.";
RL Gene 159:131-135(1995).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=98151500; PubMed=9482873;
RA Szyperski T., Fernandez C., Mumenthaler C., Wuehrlich K.;
RT "Structure comparison of human glioma pathogenesis-related protein
RT GLIPR and the plant pathogenesis-related protein Pla indicates a
RT functional link between the human immune system and a plant defense
RT system.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:2262-2266(1998).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE HUMAN BRAIN TUMOR,
CC GLOBLASTOMA MULTIFORM/ASTROCYTOMA, BUT NEITHER IN NORMAL FETAL
CC OR ADULT BRAIN TISSUE. NOR IN OTHER NERVOUS SYSTEM TUMORS.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: X91911; CAA63005.1;

```

```

DR EMBL: U16307; AAA82731.1; -.
DR HSSP: P04284; ICFE.
DR MIM: 602692; -.
DR InterPro: IPR001283; Allrgn_V5/TPX1.
DR Pfam: PF00188; SCP_1.
DR PRINTS: PR00837; V5TPX1IKE.
DR PRODOM: PD000542; Allrgn_V5/TPX1; 1.
DR SMART: SM00198; SCP_1.
DR PROSITE: PS01009; SCP_AG5_PRL_SC7_1; 1.
DR PROSITE: PS01010; SCP_AG5_PRL_SC7_2; 1.
FT CONFLICT 125 125 D->N (IN REF. 2).
FT CONFLICT 209 266 RQDQVARYSVYPGPITPRNRRTSLFLIVSVIILISV
FT TITLVOLKYPNVLVD -> DSEIYSNVTMLYIRLAHIS
T (IN REF. 2).
SQ SEQUENCE 266 AA; 30342 MW; C04378085EFB287A CRC64;

Query Match 22.88; Score 328; DB 1; Length 266;
Best Local Similarity 34.98; Pred. No. 5.2e-21;
Matches 81; Conservative 34; Mismatches 71; Indels 46; Gaps 12;

QY 9 LETTAMTLEPVLLFL-----VAGLLSPFANEDKDPATFALLTQTQVQREIVKHNELR 63
DB 1 MRTLATIAMVSVSVNSHTANILPDI-ENED-----FIKCVIHNKFR 45
QY 64 RAVSPPARMLKMEWKEAANAQKMANQCNVRS---NPKDRM-----TSLKCGENTLYMS 116
DB 46 SEVKPTASDMLYMTWDPALQIAKAMASNCQFSNTRLKPKHLPNFTSL--GENIWTG 103
QY 117 SAP-SSMSQAIOQMFDENDNDEFGVGRTPNAVYGHITQVYVYSSYLVCAGNATCPNOKY 175
DB 104 SVPIFVSALTNNYDEIDYDFKT--RICKKVGHTQVYVMAQSYKGCAGVQCP--KV 159
QY 176 LKY-----YVCOYCPAGMANRLVYVYEGAPCASCPCPN--CDGCLCTN 218
DB 160 SGFALNSGAHFICNYGPGNYPT---WPKRGATCSACPNDCIDNLICVN 208

RESULT 13
ASP_ANCCA
ID ASP_ANCCA STANDARD: PRT: 424 AA.
AC 016937;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ancylostoma secreted protein precursor.
GN ASP.
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215086; PubMed=8636085;
RA Haddon J.M., Jones B.F., Hoffman D.R., Hotez P.J.;
RT "Cloning and characterization of Ancylostoma-secreted protein. A
RT novel protein associated with the transition to parasitism by
RT infective hookworm larvae.";
RL J. Biol. Chem. 271:6672-6678(1996).
CC -1- FUNCTION: ASSOCIATED WITH THE TRANSITION TO PARASITISM BY
CC INFECTIVE HOOKWORM LARVAE.
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: U26187; AAC47001.1;
DR HSSP: P04284; ICFE.

```



```

DR InterPro: IPR001283; Allrgn_V5/Tpx1.
DR Pfam: PF00188; SCP; 1.
DR PRINTS: PR00837; V5TPXLIKE.
DR PRODOM: PD000542; Allrgn_V5/Tpx1; 2.
DR SMART: SM00198; SCP; 2.
DR PROSITE: PS01009; SCP_AGS_PRL_SC7_1; FALSE_NEG.
DR PROSITE: PS01010; SCP_AGS_PRL_SC7_2; FALSE_NEG.
KM Signal.
FT CHAIN 1 18 POTENTIAL.
FT SEQUENCE 19 424 ANCYLOSTOMA SECRETED PROTEIN.
SQ SEQUENCE 424 AA: 45735 MW: 4808295FD3035F9 CRC64;

Query Match
Best Local Similarity 31.1%; Score 274; DB 1; Length 424;
Matches 68; Conservative 37; Mismatches 70; Indels 44; Gaps 9;

OY 29 PSFPANEKDAFTALTTOVOVREIYNKNELEBRAV-----SPPARNMLKM 76
DB 218 PDVPEINOCPSNTGMTDS---VRDTFLSVHNEFRSSVARGLEPDALGNAKAKMLKM 274
OY 77 ENKKEAANAOKMANOCNVRHNSPNKDRMTSLKCGENLYMSSA-----PSMSQAIOISWPD 131
DB 275 VYDCEVEASAIRHGKNCYQSHGEDRP---GLGENITYKTVLAKDKNAKAKAQSOLMWN 331
OY 132 EYNDEFCVGPKT-----PNAVGHYTOVWVSSYLVCGNAYCPNOKVLYKYYVC 182
DB 332 ELK--EFCVGPSNVLTFTALMNRGMQIGHYIOMANDTYTKLGCAVFCNDFP----FGVC 385
OY 183 OYCPAGNANRLVPEYEGCAPCASCPCD-----NCDGLCT 217
DB 386 OYGPGENYWG--HVITYMGOPCSCPGATCSVTEGLCS 422

RESULT 14
ID VA51_VESCR STANDARD; PRT; 202 AA.
AC P35781;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Venom allergen 5.01 (Antigen 5-1) (AG5-1) (Allergen Vesp c 5.01) (Vesp
DE c V.01).
OS Vespa crabro (European hornet).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Vespoidea; Vespidae; Vespinae; Vespa.
OX NCBI_TaxID=7445;
RN [1]
RP SEQUENCE.
RC TISSUE-Venom;
RX MEDLINE=94044316; PubMed=8227862;
RA Hoffman D. R.;
RT "Allergens in Hymenoptera venom. XXV: The amino acid sequences of
RT antigen 5 molecules and the structural basis of antigenic cross-
RT reactivity."
RL J. Allergy Clin. Immunol. 92:707-716(1993).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
DR PIR: G44522; G44522.
DR PIR: G44583; G44583.
DR HSSP: P04284; ICPE.
DR InterPro: IPR001283; Allrgn_V5/Tpx1.
DR Pfam: PF00188; SCP; 1.
DR PRINTS: PR00837; V5TPXLIKE.
DR PRODOM: PD000542; Allrgn_V5/Tpx1; 1.
DR SMART: SM00198; SCP; 1.
DR PROSITE: PS01009; SCP_AGS_PRL_SC7_1; 1.
DR PROSITE: PS01010; SCP_AGS_PRL_SC7_2; 1.
KM Venom; Allergen.
FT DISULFID 4 16 BY SIMILARITY.
FT DISULFID 8 101 BY SIMILARITY.
FT DISULFID 26 94 BY SIMILARITY.
FT DISULFID 168 185 BY SIMILARITY.

```

```

SQ SEQUENCE 202 AA: 22722 MW: EDEB39905B598997 CRC64;

Query Match
Best Local Similarity 37.2%; Score 269; DB 1; Length 202;
Matches 61; Conservative 23; Mismatches 60; Indels 20; Gaps 5;

OY 45 LTTOTVOVREIYNKNELEBRAVS-----PARNMLKMEKNKEAANAOKMANO 92
DB 36 LTKQENL--ELIKQHNEFRQKVARGLETRGNPGPPAKSNMTLVNDELQIAQVMANO 93
OY 93 CNYRHSNPKDRMTSLKCGENLYMSSAPS-----SMSQAIOISWPDENDFCGVGPPTPNV 148
DB 94 CNYGHDNCRN--SAKYSVQONIAEGSTADNFSGVSNMVMKMEDEVKDYQG--SPRNKLK 151
OY 149 VGHYTOVWVSSYLVCGNAYCPNOKVLYKYYVCQCPAGNMAN 192
DB 152 VGHYTOVMATKKEIGCGSIKIENGWHRHLYVCNTPRAGNVCN 195

RESULT 15
ID VA52_VESCR STANDARD; PRT; 202 AA.
AC P35782;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Venom allergen 5.02 (Antigen 5-2) (AG5-2) (Allergen Vesp c 5.02) (Vesp
DE c V.02).
OS Vespa crabro (European hornet).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Vespoidea; Vespidae; Vespinae; Vespa.
OX NCBI_TaxID=7445;
RN [1]
RP SEQUENCE.
RC TISSUE-Venom;
RX MEDLINE=94044316; PubMed=8227862;
RA Hoffman D. R.;
RT "Allergens in Hymenoptera venom. XXV: The amino acid sequences of
RT antigen 5 molecules and the structural basis of antigenic cross-
RT reactivity."
RL J. Allergy Clin. Immunol. 92:707-716(1993).
CC -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/TPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
DR PIR: H44522; H44522.
DR PIR: H44583; H44583.
DR HSSP: P04284; ICPE.
DR InterPro: IPR001283; Allrgn_V5/Tpx1.
DR Pfam: PF00188; SCP; 1.
DR PRINTS: PR00837; V5TPXLIKE.
DR PRODOM: PD000542; Allrgn_V5/Tpx1; 1.
DR SMART: SM00198; SCP; 1.
DR PROSITE: PS01009; SCP_AGS_PRL_SC7_1; 1.
DR PROSITE: PS01010; SCP_AGS_PRL_SC7_2; 1.
KM Venom; Allergen.
FT DISULFID 4 16 BY SIMILARITY.
FT DISULFID 8 101 BY SIMILARITY.
FT DISULFID 26 94 BY SIMILARITY.
FT DISULFID 168 185 BY SIMILARITY.
SQ SEQUENCE 202 AA: 22781 MW: B710E9198413F547 CRC64;

Query Match
Best Local Similarity 36.6%; Score 265; DB 1; Length 202;
Matches 60; Conservative 24; Mismatches 60; Indels 20; Gaps 5;

OY 45 LTTOTVOVREIYNKNELEBRAVS-----PARNMLKMEKNKEAANAOKMANO 92
DB 36 LTKQENL--ELIKQHNEFRQKVARGLETRGNPGPPAKSNMTLVNDELQIAQVMANO 93
OY 93 CNYRHSNPKDRMTSLKCGENLYMSSAPS-----SWSQAIOISWPDENDFCGVGPPTPNV 148
DB 94 CNYGHDNCRN--SAKYSVQONIAEGSTADNFVSNMVMKMEDEVKDYQG--SPRNKLK 151

```

Fri Mar 14 14:00:10 2003

us-09-698-781-3.rsp

Page 10

Oy 149 VGHYTOVWYSSYLVGCGNAYCPNOKYVKYXYVCQYCPAGNMAN 192  
| | | | | : | | | : | | | | | |  
Db 152 VGHYTOVWMAKKEICGSEDYIEDGWHRRHYLVCNYPAGNVCN 195

Search completed: March 14, 2003, 05:40:55  
Job time : 21.3258 secs